Risk-taking that signals trust increases social identification

Article (Accepted Version)

RUNNING HEAD: RISK TAKING, TRUST, AND SOCIAL IDENTITY

Risk-taking that signals trust increases social identification

Tegan Cruwys*[1,2], Mark Stevens[2], Michael J. Platow[2], John Drury[3], Elyse Williams[1], Ashleigh J. Kelly[1], Margarita Weekes[1]

[1] School of Psychology, The University of Queensland, St Lucia QLD 4072, Australia.

[2] Research School of Psychology, The Australian National University, Canberra ACT 2601, Australia

[3] School of Psychology, University of Sussex, Brighton BN1 9QH, United Kingdom.

*Corresponding author information: Dr Tegan Cruwys, Research School of Psychology, The Australian National University, Canberra ACT 2601, Australia, e-mail: tegan.cruwys@anu.edu.au.

To appear in *Social Psychology*. Manuscript conditionally accepted 4th April 2020

Acknowledgements:

We thank William Anderson, Devika Ale and Rebecca Stower for their assistance as confederates in Study 2. This project was funded by the Australian Research Council, DE160100592.

Abstract

Social identification predicts many important phenomena; however, its *determinants* have received comparably little research attention. We argue that people are more likely to socially identify with others who engage in risky behavior *that implies trust* than with those who act cautiously, and test this in four experiments with over 900 participants. The experiments found support for the hypotheses across diverse risk contexts – specifically, risk of physical injury, disease risk, and financial risk. These findings indicate that others' risk taking can strengthen shared psychological group membership.

Keywords:

Social identity; group processes; risk taking; trust; solidarity

Risk-taking that signals trust increases social identification

Social identification is the sense of self-definition that we derive from psychological affiliation with others (Tajfel & Turner, 1979; Turner, Hogg, Oakes, Reicher & Wetherell, 1987). Social identification is a central concept in social psychology, with 50 years of research establishing its importance in determining attitudes, behavior, and even perception. Fields as diverse as prejudice, environmental action, and mental illness have been advanced by research examining the effects of social identity (Haslam, Reicher, & Reynolds, 2012). However, in these many thousands of research studies investigating social identification (Haslam, 2014), the vast majority have treated social identification as an independent variable, either measured or manipulated. Only a tiny fraction of these studies have explored the determinants of social identification. We argue that this represents a gap in our knowledge of social identity. Here, we propose and test whether the risk taking of others – specifically risk taking that demonstrates trust in the observer – facilitates the development of a shared social identity.

Where does shared social identity come from?

Self-categorization theory expanded upon social identity theory by specifying the contextual factors that lead a particular social identity to become salient at any given moment (Turner, Oakes, Haslam & McGarty, 1994; Oakes, Haslam & Turner, 1994). According to this framework, social identity salience is a product of: (1) a perceiver's general readiness to use a particular social categorization (e.g., Bruner, Sanford & Sherif, 1957), and (2) how good a *fit* the identity is with the current social context, both in terms of providing a meaningful point of comparison to other groups (comparative fit, e.g., Doosje et al., 1998) and in terms of consistency with one's normative expectations for group behavior (normative fit, Oakes, Turner, & Haslam, 1991). For example, a man's Canadian identity might be more likely to become salient if: (1) he has frequently thought of himself as Canadian in the past and this has been meaningful to him, and (2) he is currently in a context where his Canadian identity is relevant, such as watching a meeting between world leaders. Experimental research has shown the importance of perceiver readiness and fit in determining the psychological salience of a given social identity (e.g. Blanz, 1999). This evidence was ground-

breaking in bringing attention to specific, measurable factors in the environment that influence a person's self-perception as a group member. However, these factors have mainly been investigated in single time-point studies in the laboratory context (e.g., Platow et al., 2008), and have focused on how a pre-existing group membership comes to be salient in the moment.

Questions that have received rather less attention, however, are how groups form in the first place (the emergence of *new* social identities; Drury & Reicher, 2000), and how these groups become important and self-defining for an individual (social identification *strength*). Outside the laboratory, much of the research on the emergence of shared social identity has focused on the effect of "top down" threats to the group and intergroup conflict. For example, major threats to ingroups – whether symbolic (Riek, Mania & Gaertner, 2006) or actual (Janis, 1972; Sherif, 1966) – can increase cohesion and affiliation. Supporting these observations, experimental research has confirmed that intergroup threat increases ingroup identification (Schmid & Muldoon, 2013; for a review see Greenaway & Cruwys, 2019). However, an explicit intergroup context may not be necessary for a sense of shared identification to emerge or increase (Campbell, 1958). For example, after a major earthquake in northern Italy, children experienced an increase in subjective affiliation with other children affected by the disaster. This experience of shared identity was instrumental in facilitating supportive behavior to others following the disaster (Vezzali, Drury, Versari, & Cadamuro, 2015; see also Drury, Cocking, & Reicher, 2009). Overall, this research highlights the way in which social context can trigger the emergence and growth of shared identities, particularly when it engenders a sense of common fate (e.g., "we are all in this together").

There is also evidence of "bottom-up" processes in the formation and growth of social identity. A body of work by Postmes and colleagues has explored the formation of social identity in small groups, demonstrating that social identities are inferred and constructed through communication and interaction among individuals (Postmes, Haslam & Swaab, 2005; Postmes et al., 2006). That is, groups and social identities can form in contexts that might traditionally have been classified as *interpersonal*. Furthermore, recent studies have demonstrated that one person's degree of social identification does not develop independently of other group members. Instead, group members

converge in their degree of social identification over time, with interaction between them facilitating this process (Jans, Leach, Garcia & Postmes, 2015). When this interaction takes place in the form of coordinated behavior, such as that of choristers and dancers, people experience a subsequent increase in social identification (Koudenberg, Postmes, Gordijn & van Mourik Broekman, 2015; van Mourik Broekman, Gordijn, Koudenberg & Postmes, 2018). Taken together, these studies suggest that we *attend to cues* in the behavior of others about the presence versus the absence of shared social identity.

One cue to how other group members feel about us that has received particular attention in previous research is respect. Specifically, Tyler and Blader (2003) proposed that when a fellow ingroup member communicates respect – for example, by giving people opportunities to contribute to group decision making – this communicates shared identity and can increase those peoples' commitment to the group. The effect of respect on social identification has been demonstrated experimentally (Simon & Stürmer, 2003; Platow, Huo, Lim, Tapper, & Tyler, 2015). Research has also indicated that certain leader behaviors (e.g., the use of collective language and self-sacrificial behaviors such as accepting salary cuts) can promote increased social identification, as well as other pertinent outcomes including greater trust in the leader (in organizational, team, and small group contexts; e.g., see De Cremer & van Knippenberg, 2005; Slater et al., 2015; Stevens et al., 2019).

In this paper, we consider an intragroup process that has hitherto been unexamined as a determinant of social identity: risk taking. We argue that risk taking behaviors by others are highly relevant to social identity processes, because risk taking can allow people to communicate the *trust* they have (or do not have) in each other. Specifically, we propose that *observers attend to cues that indicate that other social actors (potential ingroup members) trust them*. Put simply, we notice the degree to which others are willing to place their fate in our hands. By willingly pursuing risky social options, people can communicate *trust* in those who control their fate.

Although risk taking has not been explored as a factor leading to social identity development, parallel investigations suggest that risk taking is linked to feelings of trust and can form part of peoples' efforts to increase social bonds (Molm, Takahashi, & Peterson, 2000). For instance, in the

context of romantic relationships, the more attractive potential partners are, the less likely it is that their behaviors such as drug use and sexual experience will be evaluated as risky (Fishbein, Hennessy, Yzer & Curtis, 2004). In other words, when people were more positively inclined towards a target, their risky behavior was seen in a more positive light. Similarly, research has indicated that a barrier to promoting safer sex practices is that unprotected sex is rarely perceived as risky by those involved, but rather seen to reflect one's trust in, and closeness to, a partner (e.g., Bell, Revino, Atkinson & Carlson, 2003; Hammer et al, 1996; Willig, 1997). Indeed, in a sample of 260 sexually transmitted disease clinic patients, Jadack et al. (1997) found that the most frequently cited reason for not using condoms was trust in one's sexual partner. Ethnographic work on needle sharing has also explored these ideas, with research showing that needle sharing primarily occurs within close social networks and is understood to be an expression of solidarity or trust (Feldman & Biernacki, 1988; Grund, Kaplan & Adriaans, 1991). Finally, there is relevant evidence that initiation and hazing activities common to some fraternities and sports teams often involve risky behaviors that may facilitate group cohesion, although this has not been investigated from a social identity perspective (Aronson & Mills, 1959; Lafferty & Wakefield, 2018).

Overall, this evidence suggests that while contextual factors are important in determining social identity, intragroup factors also contribute. In particular, we attend to information that others communicate about their degree of shared identity with us. Information about trust may be a particularly important cue to shared identity, communicating the degree to which a person is willing to take the kind of risks that would only be safe with an ingroup member (Platow et al., 2012). However, no research to date has investigated how others' risk taking behavior affects perceptions of trust and social identity. We therefore propose two hypotheses:

H1) Risk taking by an actor facilitates an increased sense of social identification with that actor.

H2) The relationship between risk taking by an actor and perceived social identification will be mediated by inferred trust (of the actor in the self).

We tested these hypotheses in four experiments. Study 1 was a vignette-based study of risk taking in a group-based scenario involving the risk of physical injury. Study 2 was a lab-based

experiment in which a confederate took a risk of disease contagion by sharing drinks with the participant. Studies 3 and 4 were pre-registered experiments in the context of financial risk taking. In Study 3 both the independent variable and the hypothesized mediator were manipulated (as has been recommended in order to establish a causal mediation chain, e.g., Jacoby & Sassenberg, 2011), while Study 4 directly compared the effect of risk taking that was explicitly attributed to trust in the observer (participant) to risk taking that was not. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in these studies below, with verbatim details of materials and manipulations available in the Supplementary Materials. Data, syntax, and pre-registrations are available at this link.

Study 1

Study 1 utilized a hypothetical vignette scenario. Specifically, we presented participants with a vignette detailing a tour group excursion, and varied only the information about the extent to which other members of the group took a risk that signalled their trust in the participant (jumping into water from an overhang after the participant had indicated it was safe to do so).

Method

Participants. Study 1 included 147 participants who were recruited via Amazon Mechanical Turk (MTurk) for a study of "First Impressions". Participants were all residents of the United States and were predominantly White (78.9%; 14.3% Black; 5.4% Hispanic/Latino/a), with a mean age of 36.86 years (*SD*=11.14, range 20-71). The sample comprised 84 Women (57.1%) and 61 Men (41.5%), and 2 who either specified their gender identity as 'Other' or for whom data were missing. Participants were diverse in their level of education, with almost half (46.9%) having completed a tertiary degree. Prior to analyses, four additional participants were excluded because they completed the experiment in more than one sitting (i.e., time taken was longer than one hour).

Procedure. Participants commenced the study by reading a vignette which described a group tour while travelling in a foreign country, in which participants had enjoyed getting to know the tour group and had finished with an excursion to a waterfall. The vignette continued: "you find an

overhang that you would like to jump into the water from. You jump and land in the water well, and call out that it's safe to jump". The two conditions then finished with one of two statements. In the *risk avoiding* condition, the vignette finished with the statement: "Several people from the tour group follow, but they gingerly climb in from the edge instead of getting a run-up like you." In the *risk taking* condition, the vignette finished with: "Several people from the tour group follow, but they jump head-first rather than feet-first like you." Power calculations indicated a sample size of 128 participants would be needed to detect η^2=0.05 with 0.8 power.

Measures.

Manipulation check. A manipulation check was included in which participants were asked "To what degree did people from the tour group take risks during the scenario?" with response options from 0=No Risks at All to 100=Very High Risks.

Social identification. Two measures of social identification were used. First, given the hypothetical nature of the scenario, four items developed specifically for this study were included to broadly assess psychological affiliation, cohesion, and willingness to interact with the group in future. These were: (1) "Do you and the tour group make a good team?" recorded on a sliding scale from 0=Absolutely Not, to 100=Absolutely; (2) "How close to the tour group would you feel in this situation?", with scale points ranging from 0=Absolutely Not Close, to 100=Completely Close; (3) "How likely is it that you and the tour group would want to talk at a later date?"; and (4) "How likely is it that you and the tour group would see each other at a later date?", each measured on separate scales with end points: 0=Absolutely Unlikely to 100=Absolutely Likely. Second, we used the Inclusion of Self in Other scale (Aron, Aron & Smollan, 1992, see also Tropp & Wright, 2001). Participants were asked to choose which set of seven versions of overlapping circles (labelled "Self" and "Tour Group") best depicted how they felt towards the tour group, from 1=not overlapping to 7=almost completely overlapping.

Perceived trust. A four-item measure of perceived group trust (α=.84) was adapted from the Partner Opinion Questionnaire (POQ; Singh et al., 2009). One item, about likelihood of exploitation in future, was excluded (from the five trust items outlined in Singh et al., 2009) due to its unsuitability

for assessing trust in the minimal one-shot interaction contexts of the present studies. The wording of the items was altered, so that the scale measured participants' perception of the *trust that tour group members felt towards him or herself*. For example, "Members of the tour group trust me not to take advantage," rated on a seven-point scale from 1=Strongly Disagree to 7=Strongly Agree.

Results

Factor analysis using principal axis factoring was conducted with the five indicators of social identification. A single factor solution was suggested both by eigenvalues (first factor=3.34, second factor=0.82) and inspection of a scree plot. This factor explained 66.78% of the variance in the five indicators.[1] A single construct was also suggested by a reliability analysis (α = .85). Therefore the factor scores from this analysis were used as our indicator of social identification in the analyses that follow.

Manipulation check. A one-way ANOVA indicated a large and significant effect on the manipulation check, such that participants perceived that substantially more risk had been taken by the tour group in the risk taking condition (head-first jumpers; *M*=77.77, *SD*=18.43) than in the risk avoiding condition (careful jumpers; *M*=32.79, *SD*=27.64), $F(1,145)=132.86$, $p<.001$, $\eta^{2=}.48$.

Evaluation of H1. A one-way ANOVA was conducted to assess whether there were significant between-group differences in social identification. These analyses found that, compared to the risk avoiding condition (*M*=-.21, *SD*=1.08), participants in the risk taking condition identified more strongly with the tour group (*M*=.23, *SD*=0.86), $F(1,145)=7.33$, $p=.008$, $\eta^{2=}.05$. H1 was therefore supported – participants felt more strongly identified with the tour group when group members took more risks, compared to those who "played it safe".

Evaluation of H2. H2 stated that the mechanism through which risk taking fosters social identification is through building a perception of trust. Participants in the risk taking condition were significantly more likely to believe that their fellow tour group members trusted them (*M*=5.14, *SD*=1.00), than participants in the risk avoiding condition (*M*=4.63, *SD*=1.14), $F(1,145)=8.69$, p .004, $\eta^{2=}.06$.

To test whether perceived group trust mediated the relationship between risk taking and social identification, a PROCESS mediation model (Hayes, 2017; model 4; 5000 bootstrapped samples) was run (see Figure 1). This analysis indicated that condition (risk avoiding versus risk taking) predicted perceived group trust, β=.24, CI: .08, .40, p=.004. Furthermore, trust predicted social identification, β=.60, CI: .47, .73, p<.001. With trust included in the model, the direct relationship between condition and social identification was not significant, β=.08, CI: -.07, .21, p=.257. Instead, the indirect effect of risk taking condition on social identification was significant, β=.14, CI: .04, .25. Results therefore supported H2 – when group members were perceived to take risks, participants were more likely to feel trusted by the group, and this in turn led to a heightened sense of social identification.

Discussion

Study 1 found that when fellow group members took risks, this led to an increase in social identification with the group. Furthermore, this effect was shown to be mediated by inferred trust – that is, group members who took more risks were perceived by participants to have placed more trust in them, and this facilitated an increase in social identification.

Of course, a substantial limitation of Study 1 was that it relied on hypothetical risk taking by a hypothetical target, rather than actual risk taking by live interaction partners. While this allowed us to exert precise control over the constructs in the interests of internal validity (Prentice & Miller, 1992), it reduced the external validity of the findings. Study 2 addresses this limitation by utilizing a confederate in a laboratory setting.

Study 2

Study 2 aimed to test H1 and H2 in a context where participants were given the opportunity to have a live interaction and form a real first impression of a stranger. Study 2 also included a baseline control condition, to investigate whether the effects on social identification were best conceptualized as an outcome of risk taking, or of risk avoidance.

Method

Participants and design. The study included 84 participants, who were recruited from a large Australian university campus to complete a "Taste Test" study. The sample comprised 64 Women (76.2%) and 20 Men (23.8%), with a mean age of 20.48 years (*SD*=5.30). The majority were first-year psychology students who received course credit for participation. Participants were randomly assigned to one of three conditions: (1) a risk taking condition, in which the confederate took a risk by willingly sharing the participants' drinks; (2) a risk avoiding condition, in which the confederate was unwilling to take the risk of sharing the participants' drinks; and (3) a control group where the confederate did not have the opportunity to engage in risk taking.

Procedure. Participants were informed that they would be completing a taste test at the same time as one other participant (when in fact, the other participant was a confederate). The experimenter stated that because the other participant (confederate) was running late, the participant should commence the taste testing component of the study. Participants were given three fruit drinks which were in their original packaging including a "sports cap" lid, and asked to try each as many times as they like and fill out a taste ratings survey. Upon the participant's completion of this task, the late participant (confederate) arrived. In the control condition, the experimenter provided the confederate with separate bottles to conduct their taste test. In the risk taking and risk avoiding conditions, the experimenter explained that, in order to save money, the confederate would be sampling from the same bottles as the participant.[2] The experimenter then pointed to a box of antibacterial wipes on the table, explaining that, if it would make the confederate more comfortable, they could use the wipes to clean the bottle top. In the risk taking condition, the confederate stated that "it's fine", and proceeded to drink directly from the bottles when completing their taste test. In the risk avoiding condition, the confederate was hesitant to share, made an obvious effort to clean the bottles with the wipes, and also poured the drink into their mouth as opposed to drinking directly from the bottles. Once the confederate finished the taste test, the participant and confederate were separated into different rooms for the participant to complete the questionnaire. The time-intensive nature of the data collection also limited our capacity to recruit a large sample. Therefore, we aimed to recruit 96 participants as our

power analysis suggested that this would be sufficient to detect a medium (η^2=0.10) effect size with 0.8 power.

Measures

Manipulation checks. The first manipulation check aimed to assess whether participants attended to the relevant information during the manipulation. This asked participants: "When you were tasting the drinks, which of the following were true?" with three response options: (1) The other participant wiped the drink bottles before drinking out of them; (2) The other participant was willing to share the same bottles that I drank out of; and (3) None of the above / Not relevant. Five additional participants answered this question incorrectly and were excluded from the sample (leaving 84 participants included in the analysis).

The second manipulation check question assessed whether participants were sensitive to the degree of risk taking that the confederate had engaged in. Participants were asked "To what degree do you think that the other participant took risks during the taste testing?" with response options from 0=No Risk Taking at All to 100=Very High Risk Taking.

Social identification. The Inclusion of Self in Other scale was retained from Study 1 as a measure of social identification. In addition, given the live interaction and interpersonal context of Study 2, five additional items from Singh et al.'s (2009) POQ (additional to those used to measure trust, see below) were added to the design as further indicators of social identification. These items were drawn from the *attraction* subscale of the POQ and assess the perceived likeability and warmth of an interaction partner whom one has just met. For example, "I would like to get to know the other participant better". Responses were measured on a seven-point scale from 1=Strongly Disagree to 7=Strongly Agree. The remainder of the POQ items were excluded due to overlap with the construct of trust, which was our theorized mediator.

Trust. As in Study 1, trust was measured with four items adapted from the POQ (Singh et al., 2009), worded such that participants were asked to infer the confederate's trust in the participant.

Results

Factor analysis with principal axis factoring was conducted with the six indicators of social identification to determine whether these items should be treated as a single construct of social identification. A single factor solution was suggested both by eigenvalues (first factor=3.59, second factor=0.83) and inspection of a scree plot. This factor explained 57.2% of the variance in the six indicators.[1] A reliability analysis also suggested a single construct adequately described the data (α = .78). Therefore, the factor scores from this analysis were used as our indicator of social identification in the analyses that follow.

Manipulation checks. An omnibus ANOVA indicated that the experimental conditions differed on the degree to which participants perceived that the confederate had taken risks, $F(1,81)=9.22$, $p<.001$, $\eta^{2=}.19$. Participants perceived that the confederate had taken significantly more risks in the risk taking condition ($M=50.58$; $SD=31.28$) than in either the risk avoiding condition ($M=28.96$; $SD=25.83$), $t(52)=-2.78$, $p=.008$, $d=.75$, or the control condition ($M=21.57$; $SD=20.37$), $t(54)=4.17$, $p<.001$, $d=1.10$. There was no significant difference in the perceived risks between the risk avoidance and control condition, $t(56)=1.22$, $p=.229$.

Evaluation of H1. A one-way ANOVA indicated significant between-group differences in social identification, $F(2,81)=3.92$, $p=.024$, $\eta^2=.09$. Participants in the risk avoiding condition ($M=-.37$, $SD=1.01$) felt significantly less identified with the confederate than participants in the risk taking condition ($M=.36$, $SD=.81$), $t(52)=-2.94$, $p=.005$, $d=0.80$. The control condition ($M=.04$, $SD=1.05$) was not significantly different from either the risk taking condition, $t(54)=1.28$, $p=.207$, or the risk avoiding condition, $t(56)=-1.52$, $p=.135$. Therefore, H1 was supported: participants reported lower social identification with strangers when they avoided the risk of sharing the participants' drinks than when they willingly shared drinks.

Evaluation of H2. The three experimental conditions differed significantly on perceived confederate trust, $F(2, 81)=5.52$, $p=.006$, $\eta^{2=}.12$. The confederate was perceived to be significantly less trusting of the participant in the risk avoiding condition ($M=3.93$, $SD=0.68$) compared to the risk taking condition ($M=4.69$, $SD=0.77$, $t(52)=-3.86$, $p<.001$, $d=1.05$) and the control condition ($M=4.51$,

SD=1.12), $t(56)=-2.35$, $p=.022$, $d=.63$. The risk taking and control conditions did not differ significantly from one another in perceived confederate trust, $t(54)=.71$, $p=.483$.

The role of perceived trust as a mediator of the relationship between risk taking and social identification was assessed using a PROCESS mediation model (Hayes, 2017; model 4; 5000 bootstrapped samples). The independent variable was specified as multicategorical with effect-coded vectors representing (1) a comparison between the risk taking and risk avoiding conditions, and (2) a comparison between the risk avoiding and control conditions (Hayes, 2017). This analysis indicated that condition (risk taking versus risk avoiding) predicted perceived confederate trust, β=.34, CI: .04, .63, $p=.027$. Furthermore, trust predicted social identification, β=.71, CI: .55, .88, $p<.001$. With trust included in the model, the relationship between condition (risk taking versus risk avoiding) and social identification was not significant, β=.11, CI: -.11, .34, $p=.324$. Instead, the indirect effect of risk taking condition on social identification through trust was significant, β=.24, CI: .05, .44.[3] Therefore, H2 was supported: when people observe a stranger taking risks (versus avoiding risks), they are more likely to infer that the stranger trusts them, and this in turn leads to a heightened sense of social identification.

Discussion

Study 2 provided further evidence that risk taking (compared to risk avoidance) by a social target promotes greater social identification with that target. This greater identification was mediated via inferred trust; participants were less likely to believe that the confederate trusted them when the confederate had avoided taking the risk to his or her own health by sharing the participant's drinks. Study 2 advances our understanding of these processes by testing the model in the context of an actual interaction between strangers, where participants believed that the risks taken (or not taken) by the confederate had real implications for that person's health. Intimate experimental settings of this nature can also confer limitations – notably a greater likelihood of demand characteristics. This possibility was somewhat mitigated in the present instance, however, by the small element of deception used (whereby participants were told that they were taking part in a taste test study). Indeed, nothing in our own observations during data collection indicated that participants had ascertained the true purpose of

the study, or the different conditions it entailed. In this regard, it is also noteworthy that the confederates were blind to the hypotheses.

Results aligned with the findings from the vignette study – that risk taking is a behavior that can convey trust in others, and thus increase social identification. The inclusion of a control group (and thus capacity to make comparisons between this and the risk taking and risk avoiding conditions) represented a further extension of Study 1. In this regard, it is noteworthy that no significant differences were observed in levels of identification and trust between participants in the risk taking and control conditions, yet significant differences were observed between participants in the risk taking and risk avoiding conditions. This illustrates that, in the context of drink sharing at least, not only does taking risks promote gains in these key outcomes, but avoiding taking a risk (that would signal trust) *reduces* both one's sense that an actor trusts them, and the extent to which one identifies with that actor.

Study 3

Study 3 aimed to test the hypotheses in a highly-powered, pre-registered experiment that manipulated not only the independent variable (risk taking of the target) but also manipulated the proposed mediator (trust). This follows recommendations that the causality of a mediation is best tested by "disabling" the hypothesized mediator under experimental conditions (Jacoby & Sassenberg, 2011; Spencer et al., 2005). Specifically, participants were recruited to participate in a modified investment game (adapted from Tanis & Postmes, 2005), which was proceeded by a messenger program in which the other player, who was actually a remote confederate, could communicate their willingness for the participant to control their financial outcomes (risk taking vs. risk avoidance) and could communicate trust in the participant explicitly (or not). As in Study 2, we also included a control condition in which the other player indicated no preference about their willingness to take risks in the investment game. H1 tested the main effect of confederate risk taking on social identification. Due to the design of Study 3, the test of H2 was slightly different from the previous studies. Specifically, we expected a main effect of confederate trust, such that participants would feel a stronger sense of social identification when the other player explicitly expressed trust in them

compared to when they did not (H2a). Furthermore, we predicted an interaction between confederate risk taking and confederate communication of trust, such that risk taking versus risk avoidance would *only* provide relevant information about social identification under conditions when trust was not explicit (H2b).

Method

Participants and design.

Participants were 386 United States resident MTurkers recruited for an "Investment Game Study". Participants were randomly assigned to one condition of a three (risk taking, risk avoidance, control) x two (explicit trust, no explicit trust) between participants design. Participants were predominantly White (78.5%; 11.4% Asian; 6.7% Black; 6.0% Hispanic/Latino/a), with a mean age of 34.41 years (*SD*=9.92, range 18-71). The sample comprised 156 Women (40.4%) and 228 Men (59.1%) and 2 who specified their gender identity as 'Other'. Participants were diverse in their level of education, with over half (54.0%) having completed a tertiary degree.

Procedure.

The experiment began with detailed instructions explaining how the investment game would work. Specifically, participants were told that they would be given 50c and would play several rounds of an investment game with another player. Participants were informed that their actual financial incentive received for the study would be their winnings from the investment game. One player was the "investor" and could choose to allocate anywhere between 0-100% of their money to the investment pool. This money would then be tripled, and the other player ("the decider") could choose to allocate the resulting money any way they wished between the two players. Participants were then given the opportunity to discuss the upcoming game with the other player in a messenger program for 60 seconds (with statements by the other player actually computer-generated and differing by condition). The timings and wording of the other player's comments were developed with reference to MTurk chat forums, and then trialed in a pilot and revised. All participants received the following message from the other player: "Hi, do you want to be the investor, or the decider?" They then

received a message in which the manipulation of risk taking was embedded, specifically: "I don't mind either way" (control conditions), "I would prefer it if I am the decider" (risk avoiding conditions), or "I'm happy for you to be the decider" (risk taking conditions). Finally, participants received a message from the other player that manipulated the expression of explicit trust, either "Ok. I trust you. Let's do it." (explicit trust), or "Ok. Let's do it." (no explicit trust). The dependent variables immediately followed the messenger program, as well as the manipulation checks and demographics. Participants were then debriefed (no investment game actually took place) and all received $1.50.

Our goal was for Study 3 to be well-powered to detect the hypothesized effects. Power calculations indicated a sample size of 378 would be needed to detect a small effect of η^2=.04 with 0.95 power). The pre-registration for Study 3 is available at this link.

Measures.

Manipulation checks. First, participants were asked "When you were talking to the other player, which of the following were true?" with response options: "The other player was happy for me to be the decider", "The other player would have preferred to have been the decider", and "The other player did not indicate any preference for who was the decider".

Second, participants were asked "Did the other player explicitly say that they trusted you to make a good decision?", with response options: "Yes, the other player said they trusted me to make a good choice" and "No, the other player did not explicitly say this". Participants who answered either of these questions incorrectly were excluded from analysis.

Measures of risk perception and trust were also included to assess whether the independent variables were effectively manipulated. As in the previous studies, risk perception was measured using the question "To what degree did the other player take risks in the investment game?" with a continuous response scale from 0 to 100. Trust was measured using the four items from the POQ as in Studies 1 and 2.

Finally, participants were asked an open-ended question about what they thought the study was about. These responses were screened to identify anyone who stated that they thought the other player was not real or that the chat program was not real.

Social identification. Three measures of social identification were included in Study 3 (a total of 12 items). These included (1) the four-item author developed measure from Study 1 (e.g., "How close do you feel to the other player?"), and (2) the Inclusion of Self in Other scale from Studies 1 and 2. A third measure was also included in Study 3, with seven items adapted from the most widely-used measures of social identification, e.g. "I am pleased to be a member of this team" (Doosje et al., 1995) and "I feel good about being a member of this team" (Ellemers et al., 1999). These specific items were chosen because of their suitability for the impression-formation context of this study.

Results

Factor analysis was conducted with principal axis factoring for the twelve indicators of social identification to determine whether these items should be treated as a single construct. Initial results indicated a two-factor solution (eigenvalues of first factor=5.87, second factor=2.05); however, items loading on the second factor were exclusively negatively-worded and so this was likely to be due to measurement variance rather than construct variance. This was further supported by a reliability analysis, which found that when the negatively-worded items were reversed, the social identification items had a high Cronbach's alpha coefficient of .86. Therefore, a single factor solution for the 12 items was specified, which explained 48.9% of the variance.[1] The factor scores from this analysis were used as our indicator of social identification in the analyses that follow.

Manipulation checks.

An additional 111 participants (22% of the original sample) were excluded due to failing manipulation checks in accordance with our pre-registered analysis plan. This was somewhat larger than anticipated, and suggested that some aspect of the design caused confusion for a minority of participants. On closer inspection, these manipulation check failures were equally likely across all conditions with the exception that participants were more likely to incorrectly infer that the

confederate had a preference in the control condition (in which the confederate indicated that they "did not mind either way" who was the decider). In some cases, these were not errors of attention, but rather that the nature of the chat program was such that participants might have indicated a preference themselves and then the confederate's comment that they were "happy either way" was understood (not unreasonably) to be acquiescence to the participant's preference. Another 18 participants were excluded because they indicated in the free-response text that they did not believe that the confederate or the chat were real. As noted above, the final sample for analysis was 386.

An omnibus one-way ANOVA indicated that the risk taking manipulation had a significant effect on participants' perceptions that the confederate had taken a risk, $F(2,383)=17.13$, $p<.001$, $\eta^2=.08$. Pairwise comparisons indicated significant differences between each of the three conditions (all $ps<.036$), such that risk taking was perceived to be highest in the risk taking condition ($M=62.66$; $SD=2.33$), moderate in the control condition ($M=50.75$; $SD=2.57$), and lowest in the risk avoiding condition ($M=43.29$; $SD=2.40$).

Similarly, a one-way ANOVA tested the relationship between the trust manipulation and participants' perceptions that the confederate trusted them. Participants felt more trusted by the other player when they explicitly stated that they trusted them ($M=5.02$; $SD=0.69$) than when no statement to this effect was made ($M=4.56$; $SD=0.82$), $F(1,384)=35.86$, $p<.001$, $\eta^2=.09$.

Evaluation of H1. To evaluate the hypotheses, a two-way ANOVA specifying main effects and the interaction of the two independent variables was conducted. This analysis indicated a significant main effect of risk taking on social identification, $F(2,380)=9.99$, $p<.001$, $\eta^2=.05$. Participants in the risk avoiding conditions ($M=-.31$, $SD=1.06$) felt significantly less identified with the confederate than participants in the risk taking conditions ($M=.12$, $SD=.95$), $t(269)=3.57$, $p<.001$, $d=0.43$. The control conditions ($M=.20$, $SD=.91$) were significantly different from the risk avoiding condition, $t(244)=4.07$, $p<.001$, $d=.52$, but not from the risk taking conditions, $t(253)=-.67$, $p=.503$. Therefore, H1 was supported: participants reported greater social identification with a stranger who was willing to take a financial risk on them compared to a stranger who was unwilling to take this risk.

Evaluation of H2. The two-way ANOVA found a significant main effect of confederate trust on social identification, $F(1,380)=15.84$, $p<.001$, $\eta^2=.04$. Supporting H2a, participants identified more strongly with the confederate when the confederate explicitly stated their trust ($M=.19$, $SD=.95$) than when they did not ($M=-.19$, $SD=1.01$), $t(384)=-3.80$, $p<.001$, $d=.39$.

Finally, the ANOVA revealed a significant two-way interaction between the manipulations of confederate risk taking and confederate trust, $F(2,380)=4.12$, $p=.017$, $\eta^2=.02$. This interaction effect is depicted in Figure 2. Supporting H2b, the difference between the risk taking and risk avoiding conditions was significant only when the confederate's trust in the participant was not explicit, $t(138)=4.17$, $p<.001$, $d=.70$. When the confederate expressed explicit trust, the social identification "penalty" for risk avoidance was no longer present, $t(129)=.85$, $p=.396$. The control condition showed a similar pattern to the risk avoiding condition (albeit leading to higher social identification on average), whereby there was a protective effect of explicit trust, $t(113)=-3.75$, $p<.001$, $d=.70$.

Discussion

Study 3 replicated the findings of Studies 1 and 2 in the context of financial investment, finding that when the target engaged in risk taking (compared to when they did not), participants perceived a stronger social identity between themselves and the target. Study 3 also provided a causal test of the hypothesized mediator of inferred trust in the target. Here, we found that the discrepancy between risk taking and risk avoidance only emerged when the degree of trust that the target placed in the participant was ambiguous, and not when the target explicitly stated that they trusted the participant. This is consistent with our theoretical model of the phenomena, in which we argue that risk taking (or risk avoidance) is a cue that communicates information about trust. However, participants no longer relied on this cue when the critical information was already available. Study 3 had several strengths, including its large sample and its pre-registered experimental design. This increased our confidence in the findings of the other two studies, including the possible effect first indicated by the results of Study 2 that avoiding taking a risk that would signal trust may reduce an observer's trust and identification at least as much as taking that risk may increase their trust and identification.

Nevertheless, a limitation of Study 3 (and Studies 1 and 2) is that it did not explicitly test whether risk taking that implies trust facilitates increased social identification to a greater extent than other kinds of risk taking that does not. That is, although findings from Studies 1-3 have provided consistent evidence for trust as a mediator in the relationship between risk taking and social identification, no direct comparisons of the effects on social identification have been made between types of risk taking that do (versus do not) signal trust in the observer. We address this in Study 4.

Study 4

Study 4 utilised the same paradigm as Study 3. Two conditions were identical to Study 3: (1) a control condition, where the confederate did not indicate a preference for investor or decider role, and (2) the explicit-trust risk taking condition, where the confederate indicated a willingness for the participant to control their financial outcomes and stated "I trust you". To these we added two further risk taking conditions. First, an *ambiguous* risk taking condition in which, similar to the no explicit trust condition in Study 3, the confederate took risks but, importantly, did not provide a clear reason why. Second, a *non-trust* risk taking condition, in which the confederate took risks but attributed this to something other than trust in the participant. H1 tested the effect of confederate risk taking which confers trust on social identification – we expected that social identification would be higher in explicit trust risk taking conditions than in all other conditions. H2 assessed whether inferred trust mediated this relationship.

Method

Participants and design.

Participants were 303 United States resident MTurkers recruited for an "Investment Game Study". Participants were randomly assigned to one of four conditions (explicit trust risk taking, ambigious risk taking ambiguous, non-trust risk taking, control) in a between participants design. Participants were predominantly White (76.9%; 9.1% Asian; 8.6% Black; 6.9% Hispanic/Latino/a), with a mean age of 34.77 years (*SD*=9.64, range 19-68). The sample comprised 102 Women (33.7%), 196 Men (64.7%), and 5 participants (1.7%) who either specified their gender identity as 'Other' or

for whom data were missing. Participants were diverse in their level of education, with 59.4% having completed a tertiary degree.

Procedure and Measures.

The procedure and measures were identical to Study 3 except the text in the chat program. All participants received the following message from the other player: "Hi, do you want to be the investor, or the decider?" They then received a message in which the manipulations of trust and risk taking were embedded, specifically: "I trust you – I'm happy for you to be the decider!" (risk taking – explicit trust); "I'm feeling lucky – I'm happy for you to be the decider!" (risk taking – ambiguous trust); "It's not enough money to worry about – I'm happy for you to be the decider!" (risk taking – non-trust); or "I don't mind either way" (control condition).

Study 4 used the same power calculation as Study 3, indicating a total sample of 252 people would be needed for a four-condition design. The pre-registration for Study 4 is available at this link.

Results

Factor analysis was conducted with principal axis factoring for the twelve indicators of social identification to determine whether these items should be treated as a single construct. Similar to Study 3, initial results indicated a two-factor solution (eigenvalues of first factor=5.20, second factor=3.16); however, items loading on the second factor were exclusively negatively-worded and so this was likely to be due to measurement variance rather than construct variance. A reliability analysis also suggested a single construct solution was adequate (α = .76). Therefore, for consistency with the other studies, a single factor solution for the 12 items was specified, which explained 43.23% of the variance. The factor scores from this analysis were used as our indicator of social identification in the analyses that follow.

Manipulation checks.

The trust manipulation check, which led to a high level of exclusions in the control condition in Study 3, was not used to exclude participants in Study 4. The other two manipulation checks were identical to Study 3, and 103 participants were excluded in accordance with our pre-registered

analysis plan. Of these, 52 respondents failed the comprehension check, including 5 who indicated that they did not believe the messenger chat was real. A further 51 respondents passed the comprehension check, but did not correctly indicate the confederate's preferred role (i.e., to be the decider or no preference). As noted above, the final sample for analysis was 303.

An omnibus one-way ANOVA indicated that the risk taking manipulation had a significant effect on participants' perceptions that the confederate had taken a risk, $F(1,299)=17.74$, $p<.001$, $\eta^2=.06$. Pairwise comparisons indicated significant differences between each of the three risk taking conditions and the control (all *ps*<.015). Risk taking was perceived to be highest in the explicit trust risk taking condition (*M*=65.94; *SD*=22.23) and ambiguous risk taking condition (*M*=60.09; *SD*=28.06), moderate in the non-trust risk taking condition (*M*=55.42; *SD*=23.32), and lowest in the control condition (*M*=44.91; *SD*=27.34).

Evaluation of H1. The explicit-trust risk taking condition was compared to all other conditions on social identification using a *t*-test. Levene's test for equality of variance was significant, $F(1,301)=4.124$, $p=0.43$, and so equal variances were not assumed. Participants reported a stronger sense of social identification when the confederate engaged in risk taking that signalled trust in the participant than in the other conditions overall, $t(126.23)=-2.00$, $p=.048$. Therefore, H1 was supported. Planned comparisons indicated that this was driven primarily by the difference between the explicit trust risk taking condition (*M*=.18; *SD*=.84) and the control condition (M = -.16; SD = 0.97, $p=.035$), with the other risk taking conditions falling in between these extremes ($M_{\text{ambiguous risk taking}}$ = .04; SD = 1.00; $M_{\text{non-trust risk taking}}$ = -.06; SD = 0.94).

Evaluation of H2. The explicit trust risk taking condition led to significantly more inferred trust, compared to the other conditions, $t(301)=-3.52$, $p<.001$. Participants felt more trusted by the other player when they explicitly stated that they trusted them than when no statement to this effect was made. Planned comparisons indicated significant differences between the explicit trust condition (*M*=5.71; *SD*=1.07) and each of the other three conditions (all *ps*<.049; $M_{\text{ambiguous risk taking}}$ = 5.36; SD = 1.03; $M_{\text{non-trust risk taking}}$ = 5.06; SD = 1.23; M_{control} = 5.07; SD = 1.07).

To test whether perceived trust mediated the relationship between risk taking and social identification, a PROCESS mediation model (Hayes, 2017; model 4; 5000 bootstrapped samples) was run. This analysis indicated that condition (explicit trust risk taking versus all other conditions) predicted perceived group trust, β=.47, p<.001. Furthermore, trust predicted social identification, β=.65, p<.001. With trust included in the model, the direct relationship between condition and social identification was not significant, β=-.05, p=.619. Instead, the indirect effect of risk taking condition on social identification was significant, β=.31, CI: .13, .49.[4] Results therefore supported H2.

Discussion

Study 4 advanced on Studies 1-3 by manipulating the reason the target gave for risk taking. Specifically, it included a condition in which the target explicitly stated that they were willing to take a risk *because* they trusted the participant, conditions in which the target gave other reasons for their willingness to engage in risk taking, and a control condition in which the target neither indicated a desire to take a risk, nor to avoid one. Through this, we were able to directly compare whether risk taking that *signals trust* facilitates greater social identification than risk taking that does not. Findings provided support for our hypotheses, such that participants in the explicit trust risk taking condition perceived a stronger social identification between themselves and the target than participants in any other condition. The effect of risk taking on social identification was also mediated by inferred trust. Taken together, findings provide support for our theoretical model of the phenomena, which suggests that when others take risks that indicate their trust in us—not just when they take any risks—we feel a stronger sense of social identification with them. As in Study 3, the large sample and pre-registered experimental design were key strengths of the study, and give us greater confidence in the robustness of the findings.

General Discussion

The four studies presented here consistently demonstrate that when a target is perceived to engage in risk taking – specifically, risk taking that indicates a degree of trust in the participant – it increases the participant's sense of a social identification with the target. Herein, risk taking was operationalized in terms of (1) jumping headfirst into an unfamiliar body of water, (2) sharing drinks,

and (3) willingness to relinquish control over a financial incentive. In Studies 1 and 2, we provided evidence that perceived trust of the target in the participant is a psychological mediator in the relationship between risk taking and social identification. In Study 3, this mediational pathway was supported experimentally, with an explicit trust statement effectively "disabling" the benefit of risk taking. In Study 4, we manipulated the reason the target gave for taking a risk, and found that participants reported higher levels of social identification when the target explicitly stated they were willing to take a risk *because they trusted* the participant, compared to when they their reasoning was ambiguous or non-trust based.

Implications

Social identification has often been measured at the individual level and treated as if it were akin to an individual difference (McGarty, 2001). We highlight, however, that social identification is actually an outcome of psychological processes that are embedded in – not independent of – one's social world. That is, social identification reflects actual variation not only in background contextual features (e.g., comparative fit; Blanz, 1999; Oakes, Turner, & Haslam, 1991), but also in the social behavior of other people. Specifically, because social identification reflects one's sense of being part of a collective, people attend to information about how *other people* feel about the group and about them. Put simply: we look for cues that social identification is, in fact, shared. A variety of behaviors might communicate social identification, such as directly expressed respect (e.g., Platow, et al., 2015). In this article, we have demonstrated the role of *risk taking* behaviors that signal trust in facilitating social identification. Indeed, our findings (particularly those from Studies 2 and 3) also indicate that when people *avoid* taking risks that would signal their trust in another person, this has a deleterious impact on the extent to which they socially identify with the person who had the opportunity to take that risk. More focused examinations of this relationship (between risk avoidance and social identification) would be a worthwhile avenue for future research.

Previously, risk taking behavior in social psychological research has typically been conceptualized as the result of social influence. That is, risk behaviors (especially substance use, e.g., Reed, Lange, Ketchie & Clapp, 2007) have often been studied as *outcomes* of an interaction between

social identification and group norms. What we have proposed is that there may be a direct relationship between risk taking behavior and social identification, independent of specific group norms. This is because, regardless of whether the risk taking is normative or counter-normative, it may communicate information about the target's perception of the observer as a "safe", trusted member of the same social group. This may help us understand why risk taking occurs in certain social contexts, such as the hazing rituals that new fraternity members are subjected to (Drout & Corsoro, 2003), or mass gatherings such as music festivals or religious ceremonies (Hopkins & Reicher, 2015).

One pending question is whether people consciously and intentionally engage in risk taking as a strategy to signal shared group membership. That is, while these experiments demonstrate that risk taking leads to the outcome of enhancing social identification, it remains for future research to explore whether building social identification is a motive that underpins risk taking. For instance, studies have shown that men are more likely to take risks in the presence of women they find attractive (Ronay & von Hippel, 2010), and the data presented here suggests that this may be strategic behavior undertaken to develop a psychological bond. However, future studies might investigate the perceived consequences of one's own risk taking, and the circumstances in which people take risks, in order to further clarify this aspect of the model.

Another issue that is critical to note is that there is every reason to expect that the relationships explored here are bidirectional. That is, just as people may be more likely to develop a strong sense of shared identity with others who take risks that imply trust, we might hypothesize that people will also be more likely to take risks with those whom they already perceive to be ingroup members. Although these hypotheses have not yet been tested directly, existing work demonstrates that shared identity leads to increased trust (Drury, Novelli, & Stott, 2015; Platow, Foddy, Yamagashi, Li & Chow, 2012) and thus an extension of this logic might be that people do not perceive risks taken with ingroup members to be as risky. Therefore, we propose that there may be a feedback loop between social identification and risk taking.

Strengths and weaknesses

Key strengths of this research program included (a) its experimental nature (which gives us confidence that the effects observed are unlikely to have been driven by confounding variables), and (b) the variety of risks that were captured across the four studies (which increases our confidence in the generalizability of the effects). It is, however, important to acknowledge that, in some instances, effect sizes were small or marginal (e.g., in the analyses pertaining to H1 in Study 4). Although this suggests that results should be generalized with some caution, it is also possible that the focal effects may be stronger in real world settings. This is because our studies represent relatively conservative tests of whether the effects are robust in highly controlled scenarios, where other plausible explanations are excluded.

The combined sample size of the four studies (N=920) could also be viewed as a strength. However, it should be noted that sample sizes for Study 2 fell slightly short of our target. Although power calculations were based on conservative estimates of potential effect sizes, further research (e.g., with a more singular focus on testing the focal relationships in a resource-intensive live interaction scenario with a large sample) would be advantageous.

A further potential limitation of this research might be that, although our theoretical model is about group processes, only Study 1 was conducted with a traditionally "group" interaction, with the other studies more interpersonal in nature. However, this methodological choice was made in order to test the hypotheses in the context of group *formation*, and specifically shed light on the process through which people evaluate and categorize novel social targets (as per self-categorization theory, Turner et al., 1987). In Study 2, for example, participants' interaction was only with one person; however, we would nevertheless argue that participants had every reason to perceive the confederate as a (potential) ingroup member – the confederate was presented as a fellow participant in the study, a fellow first year psychology student, and confederates were a similar age, social class and ethnic background. Therefore, we would argue that these findings speak to the process of social categorization and *intra*group relations, rather than (only) interpersonal impression formation.

Overall, this article has demonstrated that people are sensitive to the degree to which others take risks with them, and that this has consequences for their inferred trust and social identification.

When we observe others taking a “leap of faith” by placing enough trust in us to take risks, this makes us feel closer to them.

References

Aron, A., Aron, E., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63, 596–612.

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*, *59*(2), 177-181.

Bell, D. C., Trevino, R. A., Atkinson, J. S. & Carloson, J. W. (2003). Motivations for condom use and nonuse. *Clinical Laboratory Science, 16,* 20-33.

Blanz, M. (1999). Accessibility and fit as determinants of the salience of social categorizations. *European Journal of Social Psychology*, *29*(1), 43–74. http://doi.org/10.1002/(sici)1099-0992(199902)29:1<43::aid-ejsp913>3.0.co;2-z

Bruner, J. S., Sanford, N., & Sherif, M. (1957). On perceptual readiness. *Psychological Review*, *64*(2), 123–152.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science.*, *3*, 14–25. https://doi.org/10.1002/bs.3830030103

De Cremer, D., & Van Knippenberg, D. (2005). Cooperation as a function of leader self-sacrifice, trust, and identification. *Leadership & Organization Development Journal*, *26*(5), 355–369. https://doi.org/10.1108/01437730510607853

Doosje, B., Haslam, S. A., Spears, R., Oakes, P. J., & Koomen, W. (1998). The effect of comparative context on central tendency and variability judgements and the evaluation of group characteristics. *European Journal of Social Psychology*, *28*, 173–184.

Drout, C. E., & Corsoro, C. L. (2003). Attitudes toward fraternity hazing among fraternity members, sorority members, and non-Greek students. *Social Behavior and Personality*, *31*(6), 535-543.

Drury, J., Cocking, C., & Reicher, S. (2009). Everyone for themselves? A comparative study of crowd solidarity among emergency survivors. *British Journal of Social Psychology*, *48*(3), 487–506.

Drury, J., Novelli, D., & Stott, C. (2015). Managing to avert disaster: Explaining collective resilience at an outdoor music event. *European Journal of Social Psychology*, *45*, 533–547. http://doi.org/10.1002/ejsp.2108

Drury, J., & Reicher, S. (2000). Collective action and psychological change: The emergence of new social identities. *British Journal of Social Psychology, 39*, 579-604.

Feldman, H. W., & Biernacki, P. (1988). The ethnography of needle sharing among intravenous drug users and implications for public policies and intervention strategies. In: R .J. Battjes & R .W.

Pickens (Eds.), *Needle sharing among intravenous drug abusers: National and international perspectives* (pp. 28-39). Rockville, Md: National Institute on Drug Abuse.

Fishbein, M., Hennessy, M., Yzer, M., & Curtis, B. (2004). Romance and risk: Romantic attraction and health risks in the process of relationship formation. *Psychology, Health & Medicine*, *9*(3), 273–285. http://doi.org/10.1080/13548500410001721846

Greenaway, K. H., & Cruwys, T. (2019). The source model of group threat: Responding to internal and external threats. *American Psychologist*, *74*(2), 218–231. https://doi.org/10.1037/amp0000321

Grund, J. P. C., Kaplan, C. D., & Adriaans, N. F. P. (1991). Needle sharing in The Netherlands: An ethnographic analysis. *American Journal of Public Health*, *81*(12), 1602–1607.

Hammer, J. C., Fisher, J. D., Fitzgerald, P., & Fisher, W. a. (1996). When two heads aren't better than one: AIDS risk behavior in college-age couples. *Journal of Applied Social Psychology*, *26*(5), 375–397. http://doi.org/10.1111/j.1559-1816.1996.tb01855.x

Haslam, S. A. (2014). Making good theory practical: Five lessons for an Applied Social Identity Approach to challenges of organizational, health, and clinical psychology. *British Journal of Social Psychology*, *53*(1), 1–20. http://doi.org/10.1111/bjso.12061

Haslam, S. A., Reicher, S. D., & Reynolds, K. J. (2012). Identity, influence, and change: Rediscovering John Turner's vision for social psychology. *British Journal of Social Psychology*, *51*(2), 201–18. http://doi.org/10.1111/j.2044-8309.2011.02091.x

Hayes, A. F. (2017). Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. 2nd Edition. Guilford Press

Hock-Long, L., Henry-Moss, D., Carter, M., Hatfield-Timajchy, K., Erickson, P. I., Cassidy, A., Macauda, M., Singer, M., & Chittams, J. (2013). Condom use with serious and casual heterosexual partners: Findings from a community venue-based survey of young adults. *AIDS and Behavior*, *17*(3), 900-913. http://10.1007/s10461-012-0177-2.

Hopkins, N., & Reicher, S. (2015). The psychology of health and well-being in mass gatherings: A review and a research agenda. *Journal of Epidemiology and Global Health*. http://doi.org/10.1016/j.jegh.2015.06.001

Jacoby, J., & Sassenberg, K. (2011). Interactions do not only tell us when, but can also tell us how: Testing process hypotheses by interaction. *European Journal of Social Psychology*, *41*(2), 180–190. http://doi.org/10.1002/ejsp.762

Jadack, R. A., Fresia, A., Rompalo, A. M., & Zenilman, J. (1997). Reasons for not using condoms of clients at urban sexually transmitted diseases clinics. *Sexually Transmitted Diseases*, *24*(7), 402-408. http://doi.org/10.1097/00007435-199708000-00004

Janis, I. L. (1972). Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes. Oxford, England: Houghton Mifflin

Jans, L., Leach, C. W., Garcia, R. L., & Postmes, T. (2015). The development of group influence on in-group identification: A multilevel approach. *Group Processes & Intergroup Relations*, *18*(2), 190–209. http://doi.org/10.1177/1368430214540757

Koudenburg, N., Postmes, T., Gordijn, E. H., & van Mourik Broekman, A. (2015). Uniform and Complementary Social Interaction: Distinct Pathways to Solidarity. *Plos One*, *10*(6), e0129061. http://doi.org/10.1371/journal.pone.0129061

Lafferty, M. & Wakefield, C. (2018) Becoming part of the team: Female student athletes' engagement in initiation activities. *The Sport Psychologist*, *32*(2), 125-135.

McGarty, C. (2001). Social Identity Theory does not maintain that identification produces bias, and Self-Categorization Theory does not maintain that salience is identification: Two comments on Mummendey, Klink and Brown. *British Journal of Social Psychology*, *40*(2), 173–176. http://doi.org/10.1348/014466601164777

Molm, L. D., Takahashi, N., & Peterson, G. (2000). Risk and trust in social exchange: An experimental test of a classical proposition. *American Journal of Sociology*. https://doi.org/10.1086/210434

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). Cognition and the group: Social identity and self-categorization. In P. J. Oakes, S. A. Haslam, & J. C. Turner (Eds.), *Stereotyping and social reality* (pp. 80–103). London: Blackwell.

Oakes, P. J., Turner, J. C., & Haslam, S. A. (1991). Perceiving people as group members: The role of fit in the salience of social categorizations. *British Journal of Social Psychology*, *30*(2), 125–144.

Platow, M. J., Foddy, M., Yamagishi, T., Lim, L., & Chow, A. (2012). Two experimental tests of trust in in-group strangers: The moderating role of common knowledge of group membership. *European Journal of Social Psychology, 42*(1), 30-35.

Platow, M. J., Grace, D. M., Wilson, N., Burton, D., & Wilson, A. (2008). Psychological group memberships as outcomes of resource distributions. *European Journal of Social Psychology*, *38*, 836-851.

Platow, M. J., Huo, Y. J., Lim, L., Tapper, H., & Tyler, T. R. (2015). Social Identification Predicts Desires and Expectations for Voice. *Social Justice Research*, *28*(4), 526-549.

Postmes, T., Haslam, S. A., & Swaab, R. I. (2005). Social influence in small groups: An interactive model of social identity formation. *European Review of Social Psychology*, *16*(1), 1–42. http://doi.org/10.1080/10463280440000062

Postmes, T., Baray, G., Haslam, S. A., Morton, T. A., & Swaab, R. I. (2006). Social Identity Formation. In J. Jetten & T. Postmes (Eds.), *Individuality and the group: Advances in social identity* (p. 215). Sage Publications Limited.

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological bulletin*, *112*(1), 160-164.

Reed, M. B. B., Lange, J. E. E., Ketchie, J. M. M., & Clapp, J. D. D. (2007). The relationship between social identity, normative information, and college student drinking. *Social Influence*, *2*(4), 269–294. http://doi.org/10.1080/15534510701476617

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: a meta-analytic review. *Personality and Social Psychology Review*, *10*(4), 336–353. http://doi.org/10.1207/s15327957pspr1004_4

Ronay, R., & Hippel, W. V. (2010). The presence of an attractive woman elevates testosterone and physical risk taking in young men. *Social Psychological and Personality Science*, *1*(1), 57–64. http://doi.org/10.1177/1948550609352807

Schmid, K., & Muldoon, O. T. (2015). Perceived threat, social identification, and psychological well-being: The effects of political conflict exposure. *Political Psychology*, *36*(1), 75–92. http://doi.org/10.1111/pops.12073

Sherif, M. (1966). *In common predicament: Social psychology of intergroup conflict and cooperation*. Boston: Houghton Mifflin.

Simon, B., & Stürmer, S. (2003). Respect for group members: intragroup determinants of collective identification and group-serving behavior. *Personality and Social Psychology Bulletin*, *29*(2), 183–193. http://doi.org/10.1177/0146167202239043

Singh, R., Simons, J. J. P., Young, D. P. C. Y., Sim, B. S. X., Chai, X. T., Singh, S., & Chiou, S. Y. (2009). Trust and respect as mediators of the other- and self-profitable trait effects on interpersonal attraction. *European Journal of Social Psychology*, *39*, 1021–1038.

Slater, M. J., Barker, J. B., Coffee, P., & Jones, M. V. (2015). Leading for gold: Social identity leadership processes at the London 2012 Olympic Games. *Qualitative Research in Sport, Exercise and Health*, *7*(2), 192-209. http://dx.doi.org/10.1080/2159676X.2014.936030

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, *89*(6), 845–51. http://doi.org/10.1037/0022-3514.89.6.845

Stevens, M., Rees, T., Steffens, N. K., Haslam, S. A., Coffee, P., & Polman, R. (2019). Leaders' creation of shared identity impacts group members' effort and performance: Evidence from an exercise task. *Plos One*, *14*(7). e0218984. https://doi.org/ 10.1371/journal.pone.0218984

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worehel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey: Brooks/Cole.

Tropp, L. R., & Wright, S. C. (2001). Ingroup Identification as the Inclusion of Ingroup in the Self. *Personality and Social Psychology Bulletin*, *27*(5), 585–600. http://doi.org/10.1177/0146167201275007

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory.* Oxford: Blackwell.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*(5), 454–463.

Tyler, T. R., & Blader, S. L. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and social psychology review*, *7*(4), 349-361.

van Mourik Broekman, A., Gordijn, E. H., Koudenburg, N., & Postmes, T. (2018). Reshaping social structure through performances: Emergent solidarity between actors and observers. *Journal of Experimental Social Psychology*, *76*(June 2017), 19–32. https://doi.org/10.1016/j.jesp.2017.12.002

Vezzali, L., Drury, J., Versari, A., & Cadamuro, A. (2015). Sharing distress increases helping and contact intentions via social identification and inclusion of the other in the self: Children's prosocial behavior after an earthquake. *Group Processes & Intergroup Relations*, *19*(3), 314–327. http://doi.org/10.1177/1368430215590492

Willig, C. (1997). The limitations of trust in intimate relationships: Constructions of trust and sexual risk taking. *British Journal of Social Psychology*, *36*(2), 211-221.

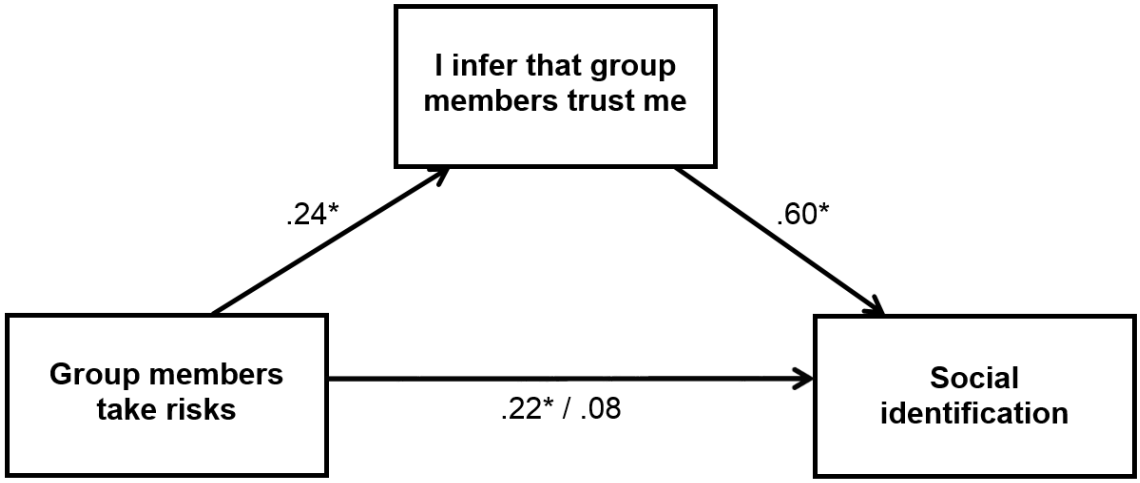

Figure 1. Study 1: Participants identify more with groups that take more risks because group members are inferred to have placed more trust in the participant.

Notes.

Standardized beta coefficients are reported.

Indirect effect=.14*

* $p < .05$.

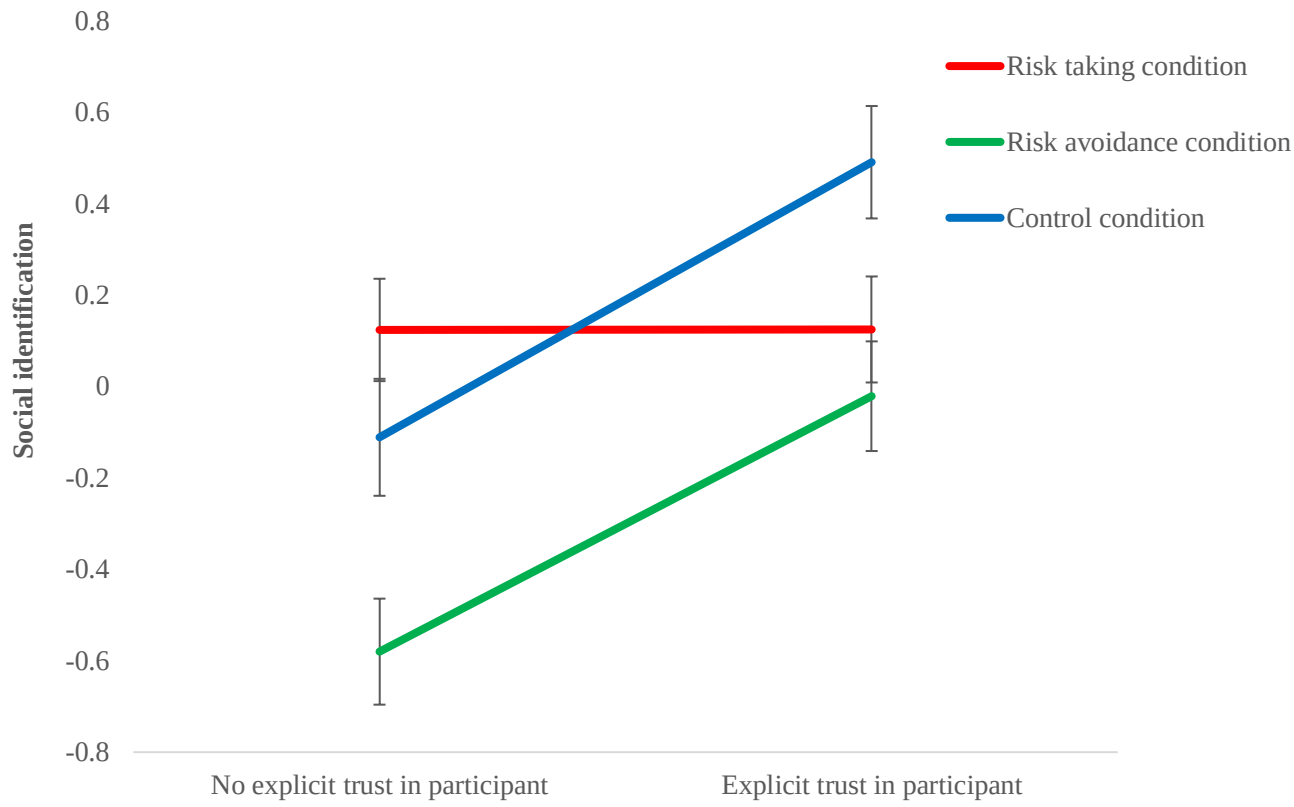

Figure 2. Study 3: Participants identify more with a stranger who takes risks, but only in the absence of an explicit assurance of that stranger's trust in them.

Note. Errors bars indicate standard errors.

Footnotes

1. In each study, the individual items comprising the full social identification scale were also analyzed separately. These analyses replicated the findings for five out of five indicators in Study 1, six out of six indicators in Study 2, and 10 out of 12 indicators in Study 3 and 12 out of 12 indicators in Study 4.
2. In fact, for the confederates' health and safety, they always drank from fresh bottles. In order to preserve the cover story for participants, the participants' bottles were surreptitiously swapped with fresh bottles that had been slightly emptied as the confederate arrived.
3. Follow-up analyses utilized a variety of different vector codes to explore all three conditions. These revealed that, in addition the risk taking versus risk avoiding effects reported in the body text, the indirect effect of experimental condition on social identity via trust was significant for the following comparisons:
    1. Risk taking versus risk-avoiding *plus* control: $\beta = .58$, CI: 28, .91
    2. Risk avoiding versus risk taking *plus* control: $\beta = .58$, CI: 29, .90.
    3. Risk avoiding versus control: $\beta = .21$, CI: .03, 40.

    However, the indirect effect was not significant for the risk taking versus control comparison. $\beta = .10$, CI: -.13, .32.
4. Follow-up analyses utilized a variety of different vector codes to explore the differences between the four conditions. These revealed that, in addition the explicit trust versus all other conditions as reported in the body text, the indirect effect of experimental condition on social identity via trust was significant for the following comparisons:
    1. Explicit trust *versus* control: $\beta = .30$, CI: .10, .58
    2. Explicit trust *versus* explicit non-trust: $\beta = .37$, CI: .14, .61.
    3. Explicit trust *versus* ambiguous trust: $\beta = .19$, CI: .01, .42.